DOCUMENT RESUME

ED 094 005 TM 003 863

AUTHOR Smith, Nick L.; Murray, Stephen L.
TITLE A Survey of Extant Development and Evaluation Models Used in Regional Educational Laboratories and Research and Development Centers.
INSTITUTION Northwest Regional Educational Lab., Portland, Oreg.
PUB DATE [Apr 74]
NOTE 17p.; Paper presented at the Annual Meeting of the American Educational Research Association (59th, Chicago, Illinois, April 1974)

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS *Evaluation Techniques; *Models; *Regional Laboratories; *Research and Development Centers; Research Problems; *Surveys

ABSTRACT

Results are presented of a survey designed to elicit information from labs and centers on the types of development models that are currently being used in product development efforts and the evaluation strategies or models that have been found to be most compatible with them. The case is made that a great deal of research needs to be conducted on evaluation and development procedures before empirically based policy can be developed. Before such a research base could be provided, two significant conceptual problems have to be addressed. First, how does one define, identify, and describe development and evaluation models per se, and secondly, how does one determine when an instructional technology perspective is indeed more relevant than an industrial management or value research perspective? (Author/RC)

ED 094005

APR 19 74

12.25

A SURVEY OF EXTANT DEVELOPMENT AND EVALUATION MODELS USED IN REGIONAL EDUCATIONAL LABORATORIES AND RESEARCH AND DEVELOPMENT CENTERS*

Nick L. Smith and Stephen L. Murray

Northwest Regional Educational Laboratory

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

TE 003 863

*Presented at the 1974 Annual Meeting of the American Educational Research Association, Chicago, April 1974.

When we proposed the topic for this paper last August, we had already spent a prestigious five weeks in the field of instructional materials development and evaluation. The survey we were proposing seemed to be a useful and necessary step in the assessment of the state of the art in instructional materials development and an interesting prologue to the two succeeding papers in today's symposium. We soon found to our chagrin, however, that such a survey could not be reasonably undertaken given the current state of the art in product development. We, therefore, turned our attention to the task of assessing just where we are currently with respect to the use of models in instructional product development and evaluation and to the task of identifying the antecedent conceptual problems which must be resolved before much productive research in the area can be undertaken. We found ourselves in the same position as Norman Hamm in his discussion of the efficacy and administration of research on the mental health of children (Hamm, 1974). He notes that his method of inquiry is "limited, if not somewhat paradoxical. What is offered is a rational analysis of our empirical enterprise as it applies to practical payoff..." p. 9. We join him in calling for more attention to a science of science including research performed on research policy.

As a first step, then, we need to examine just what is meant by the use of evaluation and development models in product development.

Purposes of Models

Individuals writing on the nature and use of models in science (c.f. Kaplan, 1964, Black, 1973) have noted several uses made of the term. Black, for example, discusses how the term "model" has been used in science to denote (1) scale models, (2) analogue models in which there is an isomorphic relationship between the model and the original entity, (3) a model which is a type of design of something worthy of imitation, and (4) theoretical models in which one uses an area better understood in order to solve a particular problem. We will argue that the term "models" as typically used in educational development and evaluation appears to refer to the third definition above, i.e., these development and evaluation models are simply designs of exemplar processes or procedures. They do not appear, for instance, to be theoretical models since they generally lack explanatory and/or predictive features. They simply appear to be examplars of desirable or commendable operating procedures and are probably best viewed as alternative, descriptive conceptualizations of development and evaluation processes. Black mentions such a use when he discusses models as a "convenience of exposition.../that/ may also help us to notice what otherwise may be overlooked, to shift the relative emphasis attached to details--in short, to see new connections." p. 496. This use of the term "models" seems particularly appropriate to the models used to guide evaluation design where there are several competing conceptualizations of appropriate evaluation procedures, each based on differing assumptions and values

concerning the evaluation process itself. Development models, however, frequently appear to be as much managerial strategies as helpful conceptual paradigms of any particular development process. The characterization of models as basic conceptual devices for heuristic use as appropriate in particular settings suggest at this point, at least, a conceptual difficulty in generalizing or being prescriptive about the use of particular models across varying circumstances and situations.

Reviews of Models

Although space does not permit a full review of all development and evaluation models currently being proffered, a few words of summary are included.

Probably the most current and comprehensive discussion of models and procedures for instructional development can be found in Baker (1973). Both Baker and Schutz (1970) have referred to two major development emphases: those having to do with a change support process, and those having to do with a product development process. They suggest that there may be differing development models and strategies appropriate to the different emphasis of these two approaches. Baker provides some reason to doubt whether there could be any generalizable development models for change support processes in pointing out that development procedures are highly idiosyncratic to particular contexts. Merrill and Boutwell (1973), in reviewing books and manuals designed to give guidance to developers of instructional systems, have noted that though they vary in terminology, the procedures recommended all basically include the following components:

behavioral objectives, pre-test, instructional activities, post-test, and revision based on empirical results.

There appears to be less uniformity in the field of evaluation models with its proliferation of models such as the CIPP Model (Stufflebeam, 1968), the Countenance Model (Stake, 1967), the Discrepancy Model (Provis, 1971), the Pathways Model (Scriven, undated), etc. Worthen and Sanders (1973) provide the most current comparison of the major evaluation models, though the collection of readings by Taylor and Cowley (1972) also provides a good overview.

In addition to the increased attention being paid by the research community to the development and use of models, there is also apparently an increasing tendency to use models as a means of specifying and operationalizing federal development and evaluation policy. Recent drafts of policy statements from the National Institute of Education (NIE) (c.f. NIE 1974a, NIE 1974b) and documents submitted to NIE for possible policy adoption, (c.f. Institute for Educational Development, 1973) suggest that NIE is taking a more prescriptive stance toward the use of evaluation and development models in its monitoring of instructional development work.

This increased prescriptiveness seems to emanate, at least in part, from NIE's response to a report to Congress made by the Comptroller General concerning a review of educational laboratory and research development center programs (Comptroller General of the United States, 1973). In its response to this report, NIE agrees, among other things, to (1) require contractors to state objectives in terms of the specific educational

changes that are expected, (2) to produce empirical evidence of the accomplishments of students who use the materials, and (3) to define more clearly the various stages of product development and testing (Comptroller General of the United States, 1973).

By NIE's own admission this increased prescriptiveness is not based on empirical evidence.

> "Generally we have lacked both the data base and the understanding of system dynamics needed for effective, rational policy making. In their absence, policies have been determined on the basis of the resolution of cross-pressures, analogies drawn from other fields, and the imagination and foresight of a few individuals." (NIE, 1973, p. 65)

Just how true this statement must be is evidenced by a brief review of the past research in the area of evaluation and development models.

Research on Models

Whether one looks at the overall rationale and empirical evidence supporting the use of development models or the justification of specific components of development procedures, the picture is equally bleak. Merrill and Boutwell (1973), after noting that almost all development procedures reviewed seemed to be based upon (a) the specification of objectives in terms of observable student behaviors, (b) the use of criterion-referenced rather than norm-referenced measurement procedures, and (c) the verification of instructional products by empirical test results, note that "the paucity of research on these premises suggest that they are considered axiomatic. Numerous propositions have been suggested, but only a handful of rather poorly executed studies have tested them." p. 95. Baker (1973), also notes that much of the literature in instructional development today prescribes

on the basis of faith alone, "legions of procedures." Noting a lack of empirical research in this area, she summarizes by saying,

> "...the field of educational development requires inquiry into the kinds of development procedures which are effective in particular contexts for various classes of organizations. From the current state of the art, one would suppose that minimally such procedures would require specification, field testing and revision as a foundation for development work. The possible ways in which each of these points might be translated into practice must be explored." p. 277.

Though the specification of behavioral objectives may be thought to be basic to most development models, the results of research on the effectiveness of using behavioral objectives as a means of improving teaching, guiding evaluation, or in the facilitation of learning, seems to be inconsistant and equivocal at best (Duchastal and Merrill, 1973; McKeachie, 1974; Stake, 1973). There are a number of conceptual and technical challenges which the criterion referenced test movement must meet before any development model could safely depend upon their wide spread development and use (Ebel, 1973; Klein and Kosecoff, 1973). Further, Baker (1973) indicates that there is not even indicative evidence that certain types of media contribute to a better learning environment than other types of media or no media at all. She suggests that developers themselves have to make evaluation studies of probably contending media to determine their relative effectiveness for the particular objectives of interest. And finally, it is becoming increasingly clear that the adoption of even the most basic "laws of learning" and hierarchial relations which has been a very significant feature of development models is largely without empirical justification as they are now conceptualized (McKeachie, 1974a, McKeachie, 1974b, White, 1973).

A look at the more global components of development models reveals the same lack of research foundation. Baker (1973) notes that while there are many guides for writing performance specifications, the research in support of such practice is not to be found. She indicates that the influence of specification on the purveyors of instruction has not been demonstrated in controlled experimental studies. Although there are many guidelines on how to write performance specifications, there is no research today to indicate the utility of such procedures to other developers. She notes that the question of the generalizability of specific revision procedures to other tasks and other subject matters have only been touched upon briefly.

Popham (1969) notes that this same situation prevails with research concerning the development of curriculum materials.

"...studies for the revision process to improve the quality of curriculum materials have not been frequently conducted. There seems to be an overriding faith in the idea that materials revised to be consistant with empirical tryout data will become better. This has not been clearly demonstrated. Certainly the manner in which revisions can be made most efficiently has not been carefully treated." p. 335.

The state of the art appears to be no better for evaluation models. Worthen (1972) has pointed to the sparcity of research on the phenomena of evaluation itself. He notes that present evaluation models "...represent (individually and collectively) an inadequate knowledge base for the guidance of educational evaluation." p. 3. He notes there is little or no information about the relative efficacy of alternative evaluation techniques and calls for research on the evaluation models themselves as well as on the critical components of most models, such as the obtrusiveness of data collection techniques on the educational phenomena being evaluated, and the

differential effectiveness of feedback provided from evaluation under differing conditions and scheduling. He also notes the need for studies to identify the most effective means of identifying goals and conducting needs assessments.

Stake and Denny (1969) have written about the lack of available techniques and knowledge base for the conduct of evaluation. They discuss, for example, the lack of techniques for presenting goals and priorities and the relative lack of good instrumentation for assessing instructional materials in classroom instruction. For example, acknowledging that there are many traditional observational schedules and techniques, they point to how little information is available on how validly these techniques can be used and how widely. In addition, they note the lack of adequate instructional assessment schedules and call for research and development which pays attention to the "...widely reported phenomenon of experimenter bias and the obtrusiveness of the classroom observer." p. 38. They also indicate a need for simple ways of meaningfully describing what the instructional program consists of and what the students do in and after training. Finally, they discuss evaluation's current lack of ability to show how students perform and to measure student's performance in evaluation-related contexts.

Thus, in evaluation as well as development, there is not only a lack of research evidence on the utility of various models or conceptual approaches, but a lack of evidence on the effectiveness of component procedures and the basic assumptions within the models themselves. It appears that, at this time, these models can not be considered anything

more than convenient conceptual mechanisms or, at most, value-based (not empirically based) claims about what constitutes "good" development and evaluation procedures.

There may be a need at the federal policy level for increased prescriptiveness of the use of evaluation and development models, but as long as there is such a lack of research evidence to support this prescriptiveness, it ought to be made clear that such recommendations are made from policymakers' past experience or value basis or from untested hypotheses concerning the appropriateness of models selected from fields other than education and not from empirical based research.

Implications for Research

The case has been made thus far that a great deal of research needs to be conducted upon evaluation and development procedures before empirically based policy can be developed. There are some basic conceptual problems, however, which must be solved before adequate research can be conducted upon the development and evaluation models themselves.

One basic problem is the conceptual delineation of just what constitutes a development or evaluation model and how one can identify and adequately describe such models. Merrill and Boutwell (1973) have discussed the lack of the systematic identification of the variables being manipulated by instructional developers and have proposed a taxonomy which would enable such identification. They note that it is often difficult for instructional material developers to describe their products unambiguously and thereby profit by the work of others in the field. In addition,

they cite the difficulties instructional researchers have in describing their experimental treatments adequately. Merrill and Boutwell have proposed a task variable taxomony as a first step in resolving some of these difficulties. Schutz (1970) notes that

> "The scope of the phenomenon, 'educational development,' has yet to be realized. While there is a vague realization that it should be possible to translate available relevant knowledge into a form that permits improved educational practice, it is erroneous to assume that the means for doing this are perfectly clear." p. 39.

Schutz further argues that the realm of educational development is still too young and rapidly growing to make it even feasible to attempt to describe the state of the art.

> "While the development domain is highly sophisticated and well differentiated in many areas, in education it has until very recently either been regarded as unnecessary or confounded with research." p. 40.

That this problem of lack of conceptual clarity is not confined to development is evidenced in the discussion by Worthen and Sanders (1973) of the continuing disagreement within the field about the most appropriate and useful definition of evaluation. (c.f. Scriven, 1972).

A second major problem to be overcome before addressing such concerns as (1) the utility of various development and evaluation models, (2) the compatibility of given evaluation and development models in a particular context, and (3) the interaction of particular evaluation and development models with the nature of the product, is the lack of conceptual clarity about the actual purpose for using development and evaluation models in materials development. An answer to the question: "What is an effective development or evaluation model?" requires that one first consider "effective for what purpose?" Models have been used frequently in the past and

apparently without a need on the part of the users for a strong empirical base. Development and evaluation models may currently be being used because of a belief (there is obviously no empirical evidence) that they will result in the efficient production of high quality products and result in efficient and relevant assessment of the worth of these products. This view of the use of models, however, is based upon an educational technology perspective and it is possible that models are used not for their utility in terms of educational technology, but because they serve as efficient political and managerial means of directing and controlling development work.

Micheal Apple (1972) has suggested that systems management procedures (in many ways similar to the current use of development and evaluation models) are actually means of increasing certainty and control:

> "It should be made clear, then, that systems approaches are not essentially neutral, nor are they performing a "scientific" function. By tending to cause its users and the other publics involved to ignore certain possible fundamental problems with schools as institutions, systems management also acts to generate and channel political sentiments supportive of the existing modes of access to knowledge and power." p. 15.

Obviously models may be chosen by researchers for their heuristic and conceptual features but used by administrators because of their monitoring and management capabilities. If one assumes that these models are being prescribed at the federal policy level, however, more for their managerial capabilities than their educational technology base, then the previous literature review is incomplete. The criteria for good development and evaluation models then lie not only in their utility in providing the most technically sound products, but in their utility for monitoring and

guiding the development process itself as well. This leads one to view the use of models as another type of management strategy and suggests research reviews and further research in the areas of industrial and educational management and not in the area of instructional technology.

The appropriate criteria and even most germane fields of research for development and evaluation models are therefore inextricably tied to the purpose or use that such models have in the instructional development process. Presumably models can fulfill multiple purposes or roles in the instructional development process, indeed Scriven (1967) sometime ago discussed the multiple roles evaluation can generally play. If, however, development and evaluation models are simply alternative heuristic conceptualizations of their respective processes, then the assessment of their validity or utility must be made in terms of the value perceptions of the relevant individuals involved, i.e., one conducts value research. If, on the other hand, these models are to be used as validated procedures for producing the technologically best instructional products, then considerably more research needs to be conducted in the field of instructional technology. Finally, if these models are to be used as a managerial control and monitoring mechanism then research in the area of industrial and educational management is most relevant. At present there appears to be little clarity in the field of instructional materials development regarding the actual purposes development and evaluation models are being used for and more importantly, the actual purposes for which such models were orginally designed.

In summary, then, it appears that there is practically no solid empirical evidence to justify the prescriptive use of development and evaluation models in instructional products development--at least from an instructional technology standpoint. Furthermore, before such a research base could be provided, two significant conceptual problems have to addressed. First, how does one define, identify and describe development and evaluation models per se, and secondly, how does one determine when an instructional technology perspective is indeed more relevant than an industrial management or value research perspective? Perhaps when these questions have been adequately answered someone will be able to reasonably conduct a survey of extant development and evaluation models in labs and centers.

References

Apple, Michael W., The adequacy of Systems Management Procedures in Education, Journal of Educational Research, 66, (1), September, 1972.

Baker, Eva L., The Technology of Instructional Development, Second Handbook of Research on Teaching, Robert M. W. Travers, (ed.) Rand McNally & Co., Chicago, 1973, pp. 245-285.

Black, Max, Models and Archetypes, Philosophy of Educational Research, H. S. Broudy, R. H. Ennis, L. I. Krimerman (eds.) John Wiley & Sons, Inc., New York, 1973.

Comptroller General of the United States, Educational Laboratory and Research and Development Center Programs Need to be Strengthened, GAO Report to the Congress, 1973.

Duchastel, D. C. and Merrill, P. F., The Effects of Behavioral Effects on Learning: A Review of Empirical Studies, Review of Educational Research, 43, (1), Winter, 1973.

Ebel, R. L., Evaluation and Educational Objectives, Journal of Educational Measurement, 1973, 10, pp. 173-279.

Hamm, Norman H., The Politics of Empiricism, American Psychologist, 29, (1), January, 1974.

Institute for Educational Development, Field Testing Education Products, Unpublished review draft, 1973.

Kaplan, Abraham, The Conduct of Inquiry, Chandler Publishing Company, San Francisco, 1964.

Klein, S. P. and Kosecoff, Jacqueline, Issues and Procedures in the Development of Criterion Referenced Tests, ERIC TM Report 26, 1973.

McKeachie, Wilbert J., Instructional Psychology, Annual Review of Psychology, 25, Annual Reviews Inc., Palo Alto, California, 1974a.

McKeachie, Wilbert J., The Decline and Fall of the Laws of Learning, Educational Researcher, 3, (3), March, 1974b.

Merrill, David M., and Boutwell, Richard C., Instructional Development: Methodology and Research, Review of Research in Education, 1, Fred N. Kerlinger (ed.) F. E. Peacock Publishers, Itasca, Illinois, 1973.

National Institute of Education, A Proposed NIE Approach to Evaluating Development Products, Unpublished review draft, 1974a.

National Institute of Education, Building Capacity for Renewal and Reform, Task Force on Resources Planning and Analysis, National Institute of Education, Washington, D. C., December, 1973.

National Institute of Education, Guidelines for Delivery of Development Products, Unpublished review draft, 1974b.

Popham, W. James, Curriculum Materials, Review of Educational Research, 39, (3), June, 1969.

Provus, Malcolm M., Discrepancy Evaluation, John McCutchan Publishing, Berekely, California, 1971.

Schutz, Richard G., The Nature of Educational Development, Journal of Research and Development in Education, 3, (2), Winter, 1970.

Scriven, Michael, Evaluation: Noble Profession and Pedestrian Practice in Curriculum, Theory Network Monograph Supplement Curriculum Evaluation: Potentiality and Reality, Joel Weiss (ed.), Ontario Institute for Studies in Education, 1972.

Scriven, Michael, The Evaluation of Education Goals, Instructional Procedures and Outcomes or the Iceman Cometh, University of California, Berekely, undated, mimeo.

Scriven, Michael, The Methodology of Evaluation, AERA Monograph Series in Curriculum Evaluation No. 1 Perspectives on Curriculum Evaluation, Robert E. Stake (ed.), Rand McNally & Company, Chicago, 1967.

Stake, Robert E., Measuring What Learners Learn, School Evaluation The Politics and Process, Ernest R. House (ed.), McCutchan Publishing Company, Berekely, California, 1973.

Stake, Robert E., Denny, Terry, Need Concepts and Techniques for Utilizing More Fully the Potential of Evaluation, NSSE 68th Yearbook, Part 2, Educational Evaluation: New Roles, New Means, National Society for the Study of Education, Chicago, 1969.

Stake, Robert E., The Countenance of Educational Evaluation, Teachers College Record, 68, (7), April, 1967.

Stufflebeam, Daniel L., Toward A Science of Educational Evaluation, Educational Technology, VIII, (14), July, 1968.

Taylor, Peter A. and Cowley, Doris M., Readings in Curriculum Evaluation, Wm. C. Brown Company Publishers, Dubuque, Iowa, 1972.

White, R. E., Research into Learning Hierarchies, Review of Educational Research, 43, (5), Summer, 1973.

Worthen, Blaine R. and Sanders, James R., Educational Evaluation: Theory and Practice, Charles A. Jones Publishing Co., Worthington, Ohio, 1973.

Worthen, Blaine R., Impediments to the Practice of Educational Evaluation, paper presented at the Annual Meeting of the American Educational Research Association, Chicago, April, 1972.